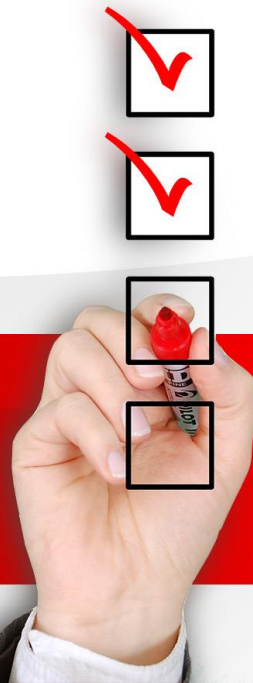


STATISTIKA 2

Pertemuan 1- REVIEW dan Sampling

Aris B. Setyawan, dan sumber relevan lainnya



Pengertian Singkat



• **Statistika Deskriptif:**

Statistika yang berfungsi untuk mendeskripsikan atau memberi gambaran terhadap obyek yang diteliti melalui data sampel atau populasi sebagaimana adanya tanpa melakukan analisis atau membuat kesimpulan yang berlaku untuk umum. Cara penyajian data biasanya berupa tabel, grafik, diagram, serta ukuran-ukuran data

• **Statistika Inferensia:**

Statistika yang berkenaan dengan penarikan kesimpulan berdasarkan data yang diperoleh dari sampel untuk menggambarkan karakteristik atau ciri dari populasi

Materi Statitika 2

1. Sampling dan Dstribusi Sampling
2. Pendugaan Parameter
3. Pengujian Hipotesis
4. Uji Chi Kuadrat
5. Analisis Varians
6. Regresi dan Korelasi
7. Statistika Non-parametrik*



Pengertian Singkat



•Populasi :

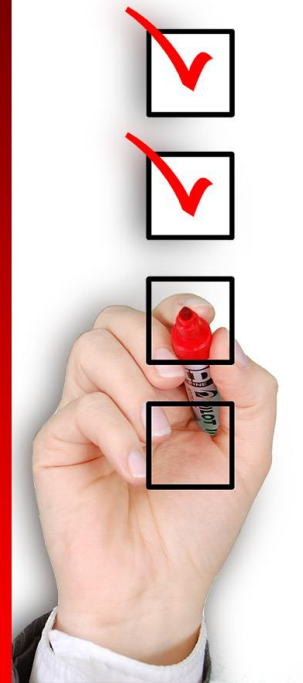
Kumpulan / keseluruhan anggota dari obyek penelitian dan memenuhi kriteria tertentu yang telah ditetapkan dalam penelitian . Penelitian yang melibatkan populasi sebagai obyek penelitian disebut Sensus

•Sampel :

Bagian tertentu dari unit populasi . Penelitian yang melibatkan sampel sebagai obyek penelitian disebut Sampling

Prinsip Dasar

1. Untuk resiko perbedaan hasil antara populasi dengan sampel, dipergunakan kemungkinan tingkat kesalahan (misalnya 1%, 5%, 10%)
2. Angka tingkat kepercayaan tersebut paralel dengan tingkat kepercayaan/ kebenaran (misalnya 99%, 95%, 90%)



Kelebihan dan Kekurangan



•Populasi :

Kelebihan :

- Data dijamin lebih lengkap
- Pengambilan kesimpulan/generalisasi lebih akurat

Kekurangan :

- Membutuh. banyak sumber daya (biaya, tenaga, waktu)
- Tidak ada jaminan bahwa semua anggota populasi dapat didata/dilacak di lapangan

•Sampel :

Kelebihan :

Efisien penggunaan sumber daya (tenaga, biaya, waktu)
Anggota sampel lebih mudah didata/dilacak di lapangan

Kekurangan :

Membutuhkan ketelitian dalam menentukan sampel
Pengambilan kesimpulan/generalisasi perlu analisis yang teliti dan dilakukan secara hati-hati

Alasan Digunakannya Sampel



1. Keterbatasan Sumber daya (Waktu, tenaga, biaya)
2. Populasi yang perlu dijaga keberadaannya (agar tidak rusak)
3. Ukuran populasi yang terlalu besar/tidak atau sulit diketahui jumlahnya dan sangat heterogen
4. Efisiensi melalui sampel yang tepat
5. Karakteristik khusus populasi (berbahaya, tersebar luas, dll)

Kriteria Sampel Yang Baik



1. Representatif (harus dapat mewakili populasi atau semua unsur sampel) → Akurat (tingkat keyakinan) dan Presisi (tingkat ketepatan)
2. Batasan sampel harus jelas
3. Dapat dilacak di lapangan
4. Tidak ada keanggotaan sampel yang ganda (di data dua kali atau lebih)
5. Harus uptodate (terbaru dan sesuai dengan keadaan saat dilakukan penelitian)

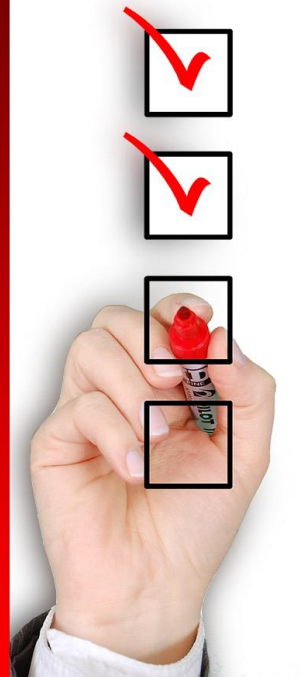
Metode Pengambilan Sample Yang Baik



1. Prosedurnya sederhana dan mudah dilakukan
2. Dapat memilih sampel yang representatif
3. Efisien dalam penggunaan sumber daya
4. Dapat memberikan informasi sebanyak-banyaknya mengenai sampel

Jumlah Sampel Yang Baik

1. Derajat keseragaman/heterogenitas dari populasi
2. Metode analisis yang akan digunakan
3. Ketersediaan sumber daya
4. Presisi yang dikehendaki



Tahapan Pemilihan Sampel



1. Penentuan Populasi : menentukan apa yang menjadi elemen populasi (individu, organisasi, produk)(individu, organisasi, produk)
2. Penentuan Unit Pemilihan Sampel : menentukan kelompok-kelompok elemen berdasarkan desain sampel yang digunakan.
3. Penentuan Kerangka Pemilihan Sampel : menentukan daftar elemen dari setiap unit pemilihan sampel.
4. Penentuan Desain Sampel : menentukan teknik sampling yang digunakan (probability sampling atau non probability sampling)
5. Penentuan Jumlah Sampel : menentukan jumlah atau besarnya sampel yang digunakan dalam penelitian
6. Pemilihan Sampel : menentukan elemen yang akan menjadi sampel dari penelitian yang dilakukan

Menentukan Ukuran Sampel



Menurut Hair et al (1998)

- Rasio antara jumlah subjek dan jumlah variabel independen dalam analisis multivariat dianjurkan sekitar 15 sampai 20 subjek per variabel independen

Menentukan ukuran sampel menurut Gay



Ukuran minimum sampel yang dapat diterima berdasarkan pada desain penelitian yang digunakan, yaitu :

- Metode deskriptif, minimal 10% populasi
- Untuk populasi yang relatif kecil min 20%
- Metode deskriptif-korelasional, minimal 30 subyek
- Metode ex post facto, minimal 15 subyek per kelompok
- Metode eksperimental, minimal 15 subyek per kelompok

Menurut GAY DAN DIEHL (1992)



- Penelitian deskriptif korelasional, paling sedikit 30 elemen populasi,
- Penelitian perbandingan kausal, 30 elemen per kelompok,
- Metode ex post facto, minimal 15 subyek per kelompok
- Penelitian eksperimen 15 elemen per kelompok.

Menurut ROSCOE (1975)



- Sebaiknya ukuran sampel di antara 30 s/d 500 elemen
- Jika sampel dipecah lagi ke dalam subsampel (laki/perempuan, SD/SLTP/SMU), jumlah minimum subsampel harus 30
- Pada penelitian multivariate (termasuk analisis regresi multivariate) ukuran sampel harus beberapa kali lebih besar (10 kali) dari jumlah variable yang akan dianalisis.
- Untuk penelitian eksperimen yang sederhana, dengan pengendalian yang ketat, ukuran sampel bisa antara 10 s/d 20 elemen.



**Penentuan Jumlah Sampel
Jika Populasi *Diketahui***

Rumus Slovin



- Rumus slovin mempersyaratkan anggota populasi diketahui jumlahnya atau populasi terhingga. Jika populasi tidak diketahui jumlah anggotanya (populasi tak terhingga), maka rumus ini tidak dapat digunakan. Terlebih lagi jika populasinya tidak diketahui keberadaan atau jumlahnya. Teknik sampling yang digunakan harus menggunakan teknik *Non Probability Sampling*. Rumus slovin ukuran sampel dengan margin eror 0,01 (1%) , 0,05 (5%) dan 0,10 (10%), yakni dengan rumus:

- $$n = \frac{N}{1 + Na^2}$$

Keterangan:

- n = Ukuran sampel yang dibutuhkan
- N = Jumlah populasi
- α = margin error yang diperkenankan atau taraf dignifikasi 0,01(1%), , 0,05 (5%) dan 0,10 (10%)
- Sebagai contoh, jika diketahui populasi penelitian sebesar 34.353 dengan taraf signifikansi 10% maka jumlah sampel yang akan dipakai dalam penelitian adalah:
- $$n = \frac{34353}{1 + 34353(0,1)^2} = 99,709 \text{ dibulatkan menjadi } 100$$

Rumus Issac dan Michael



- Untuk tingkat kesalahan 1%, 5%, dan 10%.Rumus untuk menghitung populasi sampel dari populasi yang ketahui jumlahnya adalah sebagai berikut :

$$S = \frac{X^2 NP(1-P)}{d^2(N-1) + X^2 P(1-P)}$$

$$\begin{aligned} S &= \frac{9,488 \cdot 21977 \cdot 0,5 (1-0,5)}{0,05^2 (21977-1) + 9,488 \cdot 0,5 (1-0,5)} \\ S &= 52129,44 / 54,94 + 2,372 \\ S &= 52129,44 / 57,312 \\ S &= 909,57 \end{aligned}$$

Keterangan:

- S = Ukuran sampel
- N = Ukuran populasi
- P = Proporsi dalam populasi, asumsi diambil P=0,50
- d = Ketelitian (error)
- X^2 = harga tabel chi-kuadrat untuk df tertentu
- Sebagai contoh, jika diketahui banyaknya populasi sebesar 21.977 dengan derajat ketepatan atau ketelitian sebesar 0,05 dan banyaknya variabel (df) = 4 maka banyaknya jumlah sampel yang diambil dalam penelitian adalah: 909,57 dibulatkan menjadi 910

Rumus Taro Yamane



- Keterangan:
- n = Jumlah sampel
- N = Jumlah populasi
- d = Level signifikansi yang diinginkan (umumnya 0,05 untuk bidang non eksak dan 0,01 untuk bidang eksakta.
- Sebagai contoh, jika diketahui populasi penelitian sebesar 21.977 dengan taraf signifikansi 5% maka jumlah sampel yang akan dipakai dalam penelitian adalah

$$n = \frac{N}{Nd^2 + 1}$$

$$n = \frac{21977}{21977(0,05)^2 + 1} = 392,84 \text{ atau dibulatkan menjadi } 393$$

Rumus Rao Purba



populasi yang tak hingga. Menurut Rao Purba pada prinsipnya tidak ada aturan yang pasti untuk menentukan presentasi yang dianggap tepat dalam menentukan sampel. *Margin of error maximum* atau kesalahan maksimal yang bisa diterima ditetapkan dalam rumus ini sebesar 0,10 atau 10%.

- Keterangan:
- n
- N = Jumlah populasi
- Moe = *Margin of error maximum*
- Sebagai contoh, jika diketahui populasi penelitian sebesar 21.977 dengan *margin of error maximum* sebesar 10% maka jumlah sampel yang akan dipakai dalam penelitian adalah :

$$n = \frac{N}{(1 + N \cdot Moe^2)}$$

$$n = \frac{21977}{(1 + 21977 \cdot 0,1^2)} = 99,547 \text{ atau dibulatkan menjadi } 100.$$

Populasi (N)	Sampel (n)	Populasi (N)	Sampel (n)	Populasi (N)	Sampel (n)
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346



Menurut Krejcie dan Morgan (1970)

Populasi (N)	Sampel (n)	Populasi (N)	Sampel (n)	Populasi (N)	Sampel (n)
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000	384



Lanjutan....Krejcie dan Morgan
(1970)



Penentuan Jumlah Sampel Jika Populasi *Tidak* *Diketahui*

Rumus Jacob Chohen



- Jacob Cohen (1988) menetapkan ukuran sampel berdasarkan teknik analisis datanya. Ada empat faktor yang perlu dilihat dalam penentuan ukuran sampel agar dapat memenuhi *statistic power analysis* yaitu *sample size, signficancy, directionality and effect size*. Berikut adalah rumus Jacob Chohen

- $$N = \frac{t}{f^2} + u + 1$$

- Keterangan:
- N = Ukuran sampel(*sample size*)
- f^2 = *Effect Size* (umumnya sebesar 10%)
- u = Banyaknya variabel yang terkait dalam penelitian
- t = t tabel diperoleh dari tabel disignifikasi 1% dengan banyaknya variabel dalam penelitian (u)

Sebagai contoh

- Sebagai contoh, jika diketahui $p = 0,95$ dan effect size (f^2) = 0,1 dengan terdapat 5 variabel yang terkait dalam penelitian (u). Nilai t tabel dengan taraf signifikan 1% dan $p = 0,95$ dan $u = 5$ adalah 19,76.
- $$N = \frac{19,76}{0,1} + 5 + 1 = 203,6 \text{ dibulatkan menjadi } 204$$

Rumus Paul Leedy



- Model Paul Leedy digunakan jika populasi penelitian merupakan sebuah proporsi atau bagian dari kelompok populasi lain yang lebih besar ukurannya, seperti misalnya populasi berupa jumlah keluarga miskin disuatu daerah, dimana keluarga miskin merupakan bagian dari jumlah KK yang ada di daerah tersebut

Keterangan :

- N = Ukuran sampel
- Z = Standard score untuk α yang dipilih
- e = Sampling error yang dipergunakan
- P = Proporsi harus dalam populasi
- Sebagai contoh, jika jumlah populasi dari suatu penelitian tidak diketahui maka harga P (1-P) maksimal adalah 0,25 dan menggunakan *Confidence Level* 95% dengan tingkat kesalahan tidak lebih dari 10%, maka besar sampel adalah
- $$N = \left(\frac{1,96}{0,1} \right) (0,5) \left(1 - \frac{1}{0,5} \right)$$
- N = 96,04 dibulatkan menjadi 100

$$N = \left(\frac{Z^2}{e} \right) (P)(1 - P)$$